BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: June 17, 2003, 11:16:03 ; Search time 236.058 Seconds
(without alignments)
10331.847 Million cell updates/sec

Title: US-09-807-933B-6
Perfect score: 1083
Sequence: 1 atgaagttccttaccattgc.....ctggctgttcaagaaaataa 1083

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1083	100.0	1083	21	AAA62728	Endoglucanase nucl
2	1083	100.0	1083	24	AAL43246	Rhizopus arrhizus
3	519.2	47.9	1017	21	AAA62726	Endoglucanase nucl
4	519.2	47.9	1017	24	AAL43244	Rhizopus arrhizus
5	519	47.9	1101	21	AAA62727	Endoglucanase nucl
6	519	47.9	1101	24	AAL43245	Rhizopus arrhizus
7	404.6	37.4	1017	21	AAA62729	Endoglucanase nucl
8	404.6	37.4	1017	24	AAL43247	Rhizopus arrhizus
9	398.6	36.8	1164	21	AAA62730	Endoglucanase nucl
10	398.6	36.8	1164	24	AAL43248	Rhizopus arrhizus
11	363.4	33.6	1041	21	AAA62731	Endoglucanase nucl
12	363.4	33.6	1041	24	AAL43249	Phycomyces nitens
13	362.4	33.5	1043	21	AAA62732	Endoglucanase nucl
14	362.4	33.5	1043	24	AAL43250	Rhizopus arrhizus
15	218.2	20.1	1473	12	AAQ14857	Fusarium oxysporum
16	218.2	20.1	1473	13	AAQ26407	Fusarium oxysporum
17	218.2	20.1	1473	13	AAQ26382	Endoglucanase #2.
18	218.2	20.1	1473	13	AAQ25933	Cellulase containe
19	218.2	20.1	1473	13	AAQ29935	Endoglucanase gene
20	218.2	20.1	1473	14	AAQ49942	Endoglucanase enzy
21	218.2	20.1	1473	16	AAZ60179	F. oxysporum endog
22	218.2	20.1	1473	19	AAV16103	Fusarium oxysporum
23	216.6	20.0	1473	14	AAQ41733	Dye transfer inhib
24	215.6	19.9	984	19	AAV16105	Fusarium oxysporum
25	205.4	19.0	1423	17	AAT39049	cDNA encoding cell
26	195	18.0	928	19	AAV15074	Hybrid DNA compris
27	184.4	17.0	922	19	AAV15072	Hybrid DNA compris
28	184.4	17.0	922	19	AAV15073	Hybrid DNA compris
29	181	16.7	1174	17	AAT39050	cDNA encoding cell
30	181	16.7	1174	19	AAV39096	Monocomponent endo
31	178.2	16.5	927	17	AAT39062	Chimeric endogluca
32	178	16.4	960	17	AAT39047	cDNA encoding cell
33	177	16.3	894	17	AAT39061	Chimeric endogluca
34	171.8	15.9	1261	19	AAV23748	Humicola grisea ce
35	170.2	15.7	913	17	AAT39051	cDNA encoding cell
36	168.4	15.5	885	17	AAT39075	Chimeric endogluca
37	162.4	15.0	1060	13	AAQ30072	43kD endoglucanase
38	162	15.0	672	24	AAL43263	Humicola insolens
39	162	15.0	672	24	AAI69425	Humicola insolens
40	161.8	14.9	1058	13	AAQ26405	Humicola insolens
41	161.8	14.9	1060	12	AAQ14856	Humicola insolens
42	161.8	14.9	1060	13	AAQ26380	Endoglucanase #1.
43	161.8	14.9	1060	13	AAQ25932	Cellulase containe
44	161.8	14.9	1060	13	AAQ29934	Endoglucanase gene
45	161.8	14.9	1060	13	AAQ30067	Sequence encoding

ALIGNMENTS

RESULT 1
AAA62728
ID AAA62728 standard; DNA; 1083 BP.
XX
AC AAA62728;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 3.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09823.
XX
PT Endoglucanases of fungal origin with high activity under alkaline

PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 113-115; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

Query Match 100.0%; Score 1083; DB 21; Length 1083;
Best Local Similarity 100.0%; Pred. No. 5.6e-302;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60
Db    1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60

Qy   61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
Db   61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Qy  121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
Db  121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Qy  181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240
Db  181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Qy  241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300
Db  241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Qy  301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360
Db  301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Qy  361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420
Db  361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Qy  421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480
Db  421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Qy  481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540
Db  481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Qy  541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600
Db  541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600

Qy  601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660
Db  601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660

Qy  661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720
Db  661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720

Qy  721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780
Db  721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780

Qy  781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840
Db  781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840

Qy  841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900
Db  841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900

Qy  901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960
Db  901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960

Qy  961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020
Db  961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020

Qy 1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080
Db 1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080

Qy 1081 TAA 1083
Db 1081 TAA 1083
```

RESULT 2
AAL43246
ID AAL43246 standard; DNA; 1083 BP.
XX
AC AAL43246;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 3.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15054.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 65-68; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1083 BP; 260 A; 297 C; 231 G; 295 T; 0 other;

```
  Query Match           100.0%;  Score 1083;  DB 24;  Length 1083;
  Best Local Similarity  100.0%;  Pred. No. 5.6e-302;
  Matches 1083;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Qy        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Qy        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Qy        541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600

Qy        601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660

Qy        661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720

Qy        721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780

Qy        781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840

Qy        841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900

Qy        901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960

Qy        961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020

Qy       1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080

Qy       1081 TAA 1083
              |||
Db       1081 TAA 1083
```

RESULT 3
AAA62726
ID AAA62726 standard; DNA; 1017 BP.
XX
AC AAA62726;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 1.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09821.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 104-105; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

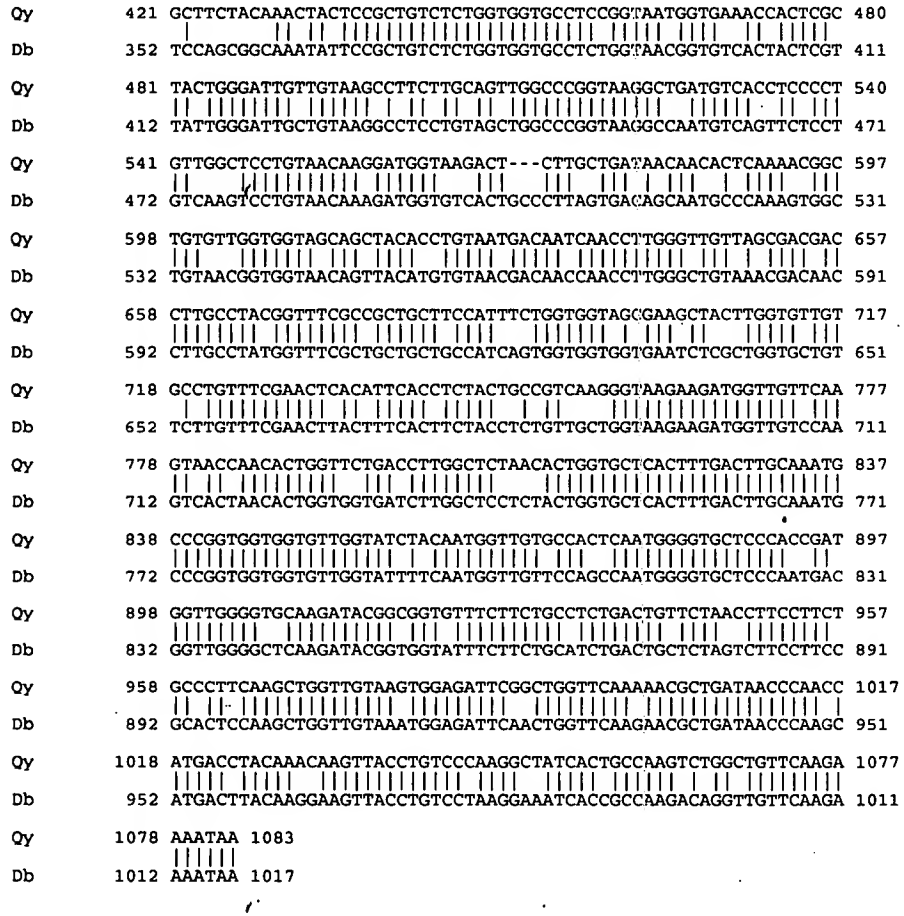
```
  Query Match           47.9%;  Score 519.2;  DB 21;  Length 1017;
  Best Local Similarity  71.9%;  Pred. No. 2.3e-139;
  Matches  781;  Conservative    0;  Mismatches  233;  Indels   72;  Gaps    5;

Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60
              ||||||||  |||| ||||||| |||||| |||||||  || ||| ||||||||||||||
Db          1 ATGAAGTTTATTACTATTGCCTCTTCCGCTCTCTTGGCTCTCGCCCTCGGTACTGAAATG 60

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
              |||  ||||||||||||||||||    ||    |||||||||||||||||||| ||||
Db         61 GCCTCTGCTGCTGAATGTAGCAAATTGTATGGTCAATGTGGTGGTAAGAACTGGAATGGC 120

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
              ||||| || |||||||||| || || ||      | |      |||   || ||||||
```

```
Db        121 CCTACTTGTTGTGAATCTGGATCCACCTG------TAAAGTAAGCAACGATTACTACTCT 174

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Db        175 CAATGTCTTCCCTCTGGAAGCAGTGGC----------------------AATAAATCT 210

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Db        211 TCTGAAAGTGCTCACAAGAAGACTACCACTGCTGCTCACAAGAAG--------------- 255

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Db        256 ------------------ACTACTACCGCTGCTCATAAAAAGACTACCACTGCTCCTGCT 297

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Db        298 AAGAAGACTACAACTGTTGCCAA------AGCTTCCACCCCTTCTAACTCTAGCTCTAGC 351

Qy        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Db        352 TCCAGCGGCAAATATTCCGCTGTCTCTGGTGGTGCCTCTGGTAACGGTGTCACTACTCGT 411

Qy        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Db        412 TATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGTAAGGCCAATGTCAGTTCTCCT 471

Qy        541 GTTGGCTCCTGTAACAAGGATGGTAAGACT---CTTGCTGATAACAACACTCAAAACGGC 597

Db        472 GTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGTGACAGCAATGCCCAAAGTGGC 531

Qy        598 TGTGTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGAC 657

Db        532 TGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAACCTTGGGCTGTAAACGACAAC 591

Qy        658 CTTGCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGT 717

Db        592 CTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGTGGTGAATCTCGCTGGTGCTGT 651

Qy        718 GCCTGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAA 777

Db        652 TCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCTGGTAAGAAGATGGTTGTCCAA 711

Qy        778 GTAACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATG 837

Db        712 GTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGTGCTCACTTTGACTTGCAAATG 771

Qy        838 CCCGGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGAT 897

Db        772 CCCGGTGGTGGTGTTGGTATTTTCAATGGTTGTTCCAGCCAATGGGGTGCTCCCAATGAC 831

Qy        898 GGTTGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCT 957

Db        832 GGTTGGGGCTCAAGATACGGTGGTATTTCTTCTGCATCTGACTGCTCTAGTCTTCCTTCC 891

Qy        958 GCCCTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACC 1017

Db        892 GCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACGCTGATAACCCAAGC 951

Qy       1018 ATGACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGA 1077

Db        952 ATGACTTACAAGGAAGTTACCTGTCCTAAGGAAATCACCGCCAAGACAGGTTGTTCAAGA 1011

Qy       1078 AAATAA 1083

Db       1012 AAATAA 1017
```

RESULT 4
AAL43244
ID AAL43244 standard; DNA; 1017 BP.
XX
AC AAL43244;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 1.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15052.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Example 10; Page 56-58; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1017 BP; 240 A; 250 C; 235 G; 292 T; 0 other;

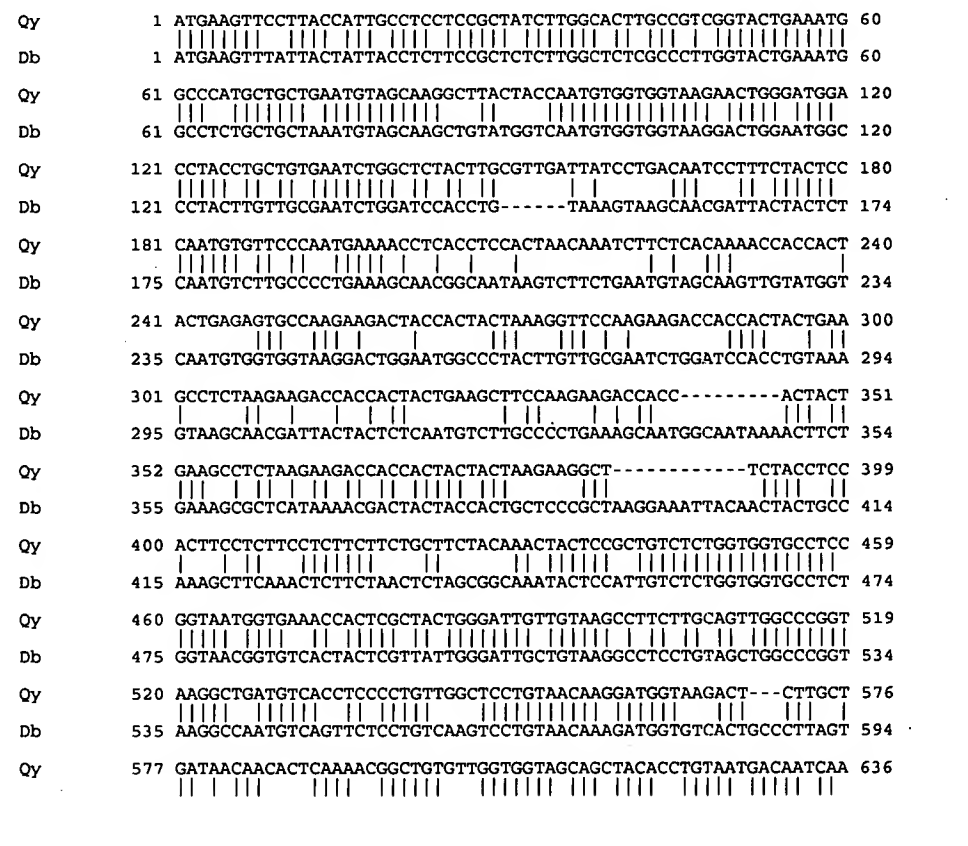
Query Match 47.9%; Score 519.2; DB 24; Length 1017;
Best Local Similarity 71.9%; Pred. No. 2.3e-139;
Matches 781; Conservative 0; Mismatches 233; Indels 72; Gaps 5;

```
Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60

Db          1 ATGAAGTTTATTACTATTGCCTCTTCCGCTCTCTTGGCTCTCGCCCTCGGTACTGAAATG 60

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Db         61 GCCTCTGCTGCTGAATGTAGCAAATTGTATGGTCAATGTGGTGGTAAGAACTGGAATGGC 120

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Db        121 CCTACTTGTTGTGAATCTGGATCCACCTG------TAAAGTAAGCAACGATTACTACTCT 174

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Db        175 CAATGTCTTCCCTCTGGAAGCAGTGGC----------------------AATAAATCT 210

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Db        211 TCTGAAAGTGCTCACAAGAAGACTACCACTGCTGCTCACAAGAAG--------------- 255

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Db        256 ------------------ACTACTACCGCTGCTCATAAAAAGACTACCACTGCTCCTGCT 297

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Db        298 AAGAAGACTACAACTGTTGCCAA------AGCTTCCACCCCTTCTAACTCTAGCTCTAGC 351
```



RESULT 5
AAA62727
ID AAA62727 standard; DNA; 1101 BP.
XX
AC AAA62727;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 2.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09822.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 108-110; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;

Query Match 47.9%; Score 519; DB 21; Length 1101;
Best Local Similarity 69.7%; Pred. No. 2.7e-139;
Matches 772; Conservative 0; Mismatches 305; Indels 30; Gaps 4;



```
Db        595 GACAGCAATGTCCAAAGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAG 654

Qy        637 CCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGT 696
              ||||||| ||| | ||| | |||||||| |||||||| |||||| |||| ||||||||||
Db        655 CCTTGGGCTGTAAACGATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGT 714

Qy        697 AGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCACCTCTACTGCCGTCAAG 756
              | ||| ||   ||||| ||| | |||||||||||| || ||||| |||||  | ||
Db        715 GGTGAATCTCGCTGGTGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCT 774

Qy        757 GGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCTTGGCTCTAACACTGGT 816
              |||||||||||||||| | |||||| || ||||||||  ||| ||||||||     ||||||
Db        775 GGTAAGAAGATGGTTATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGT 834

Qy        817 GCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACT 876
              ||||||||||||||||||||||||||||||||||||||||||| | |||||||||  |||
Db        835 GCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAG 894

Qy        877 CAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCT 936
              ||||||||||||||||  || ||||||||  | |||||||| ||| ||||||||||| |||
Db        895 CAATGGGGTGCTCCCAATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCT 954

Qy        937 GACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTC 996
              ||||| ||||  |||||||| || || |||||||||||||| |||||||||  |||||||
Db        955 GACTGCTCTAGTCTTCCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTC 1014

Qy        997 AAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTGTCCCAAGGCTATCACT 1056
              || ||||||||||||||| |||||| |||||  ||||||||||||||||  |||||
Db       1015 AAGAACGCTGATAACCCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACC 1074

Qy       1057 GCCAAGTCTGGCTGTTCAAGAAAATAA 1083
              |||||| | || ||||||||||||||||
Db       1075 GCCAAGACAGGTTGTTCAAGAAAATAA 1101
```

```
RESULT 6
AAL43245
ID   AAL43245 standard; DNA; 1101 BP.
XX
AC   AAL43245;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 2.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Rhizopus arrhizus.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15053.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 60-63; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1101 BP; 268 A; 258 C; 257 G; 318 T; 0 other;
```

```
  Query Match           47.9%;  Score 519;  DB 24;  Length 1101;
  Best Local Similarity  69.7%;  Pred. No. 2.7e-139;
  Matches  772;  Conservative    0;  Mismatches  305;  Indels   30;  Gaps    4;

Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60
              |||||||  |||| ||| |||| |||||| ||||||| || ||| | ||||||||||||
Db          1 ATGAAGTTTATTACTATTACCTCTTCCGCTCTCTTGGCTCTCGCCCTTGGTACTGAAATG 60

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
              |||  ||||||| ||||||||||   ||  ||||||||||||||| |||| ||||
Db         61 GCCTCTGCTGCTAAATGTAGCAAGCTGTATGGTCAATGTGGTGGTAAGGACTGGAATGGC 120

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
              ||||| || |||||||| || || ||      | |       ||| || ||||||
Db        121 CCTACTTGTTGCGAATCTGGATCCACCTG------TAAAGTAAGCAACGATTACTACTCT 174

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240
              |||||| || ||  ||||| |  |  |  |           | |  |||        |
Db        175 CAATGTCTTGCCCCTGAAAGCAACGGCAATAAGTCTTCTGAATGTAGCAAGTTGTATGGT 234

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300
              |||  |||  |     |     |||   ||| | |          ||||   | ||
Db        235 CAATGTGGTGGTAAGGACTGGAATGGCCCTACTTGTTGCGAATCTGGATCCACCTGTAAA 294

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACC---------ACTACT 351
              |    ||  |   |  | ||        | ||    | | ||           ||| ||
Db        295 GTAAGCAACGATTACTACTCTCAATGTCTTGCCCCTGAAAGCAATGGCAATAAAACTTCT 354

Qy        352 GAAGCCTCTAAGAAGACCACCACTACTACTAAGAAGGCT------------TCTACCTCC 399
              |||  | || | || || || ||||| |||       |||             |||| ||
Db        355 GAAAGCGCTCATAAAACGACTACTACCACTGCTCCCGCTAAGGAAATTACAACTACTGCC 414

Qy        400 ACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCC 459
              |  | ||  |||||||     ||      || ||||||  ||||||||||||||||||
Db        415 AAAGCTTCAAACTCTTCTAACTCTAGCGGCAAATACTCCATTGTCTCTGGTGGTGCCTCT 474

Qy        460 GGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGT 519
              ||||| ||||  || ||||| || |||||||| |||||| | || || || |||||||||
Db        475 GGTAACGGTGTCACTACTCGTTATTGGGATTGCTGTAAGGCCTCCTGTAGCTGGCCCGGT 534

Qy        520 AAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGATGGTAAGACT---CTTGCT 576
              ||||||  ||||||  || |||||     |||||||||| ||||||   |||  |||  |
Db        535 AAGGCCAATGTCAGTTCTCCTGTCAAGTCCTGTAACAAAGATGGTGTCACTGCCCTTAGT 594

Qy        577 GATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAGCTACACCTGTAATGACAATCAA 636
              || | |||     ||||  |||||||  |||||||| ||| |||| |||||  ||||| ||
Db        595 GACAGCAATGTCCAAAGTGGCTGTAACGGTGGTAACAGTTACATGTGTAACGACAACCAG 654

Qy        637 CCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGT 696
              ||||||| ||| | |||  | |||||||| |||||||| |||||| ||||   |||||||
Db        655 CCTTGGGCTGTAAACGATAATCTTGCCTATGGTTTCGCTGCTGCTGCCATCAGTGGTGGT 714

Qy        697 AGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCACCTCTACTGCCGTCAAG 756
              | ||| ||   ||||| ||| | |||||||||||| || ||||| |||||  | ||
Db        715 GGTGAATCTCGCTGGTGCTGTTCTTGTTTCGAACTTACTTTCACTTCTACCTCTGTTGCT 774

Qy        757 GGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCTTGGCTCTAACACTGGT 816
              |||||||||||||||| | |||||| || ||||||||  ||| ||||||||     ||||||
Db        775 GGTAAGAAGATGGTTATCCAAGTCACTAACACTGGTGGTGATCTTGGCTCCTCTACTGGT 834

Qy        817 GCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACT 876
              ||||||||||||||||||||||||||||||||||||||||||| | |||||||||  |||
Db        835 GCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATTTTCAATGGTTGCTCCAAG 894
```

```
Qy        877 CAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCT 936
              ||||||||||||||||  || ||||||||  | ||||||||  ||| ||||||||||  |||
Db        895 CAATGGGGTGCTCCCAATGACGGTTGGGGCTCGAGATACGGTGGTATTTCTTCTGCATCT 954

Qy        937 GACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTC 996
              |||||  ||||  ||||||||  ||  ||  ||||||||||||||  |||||||||  |||||||
Db        955 GACTGCTCTAGTCTTCCTTCCGCACTCCAAGCTGGTTGTAAATGGAGATTCAACTGGTTC 1014

Qy        997 AAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTGTCCCAAGGCTATCACT 1056
              ||  ||||||||||||||||  ||||||  |||||  ||||||||||||||||||  |||||
Db       1015 AAGAACGCTGATAACCCAAGCATGACTTACAAGGAAGTTACCTGTCCCAAGGAAATCACC 1074

Qy       1057 GCCAAGTCTGGCTGTTCAAGAAAATAA 1083
              ||||||  |  ||  ||||||||||||||||
Db       1075 GCCAAGACAGGTTGTTCAAGAAAATAA 1101


RESULT 7
AAA62729
ID   AAA62729 standard; DNA; 1017 BP.
XX
AC   AAA62729;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 4.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Mucor circinelloides.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09824.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs -
XX
PS   Claim 44; Page 118-119; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;

  Query Match             37.4%;  Score 404.6;  DB 21;  Length 1017;
  Best Local Similarity   65.6%;  Pred. No. 2.6e-106;
  Matches  671;  Conservative    0;  Mismatches  289;  Indels   63;  Gaps    3;

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
              ||  | ||||||   || |||   |  ||   ||||||||||| |     |||  ||||
Db         58 GCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGA 117

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
              |||||||| ||||||  ||||||||||||||||| | |    | |||    | ||||||
Db        118 CCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCT 177

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240
              ||||||  ||||          |    ||  || ||||||  ||  | ||   | || |||
Db        178 CAATGTCTTCC---------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACC 228

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300
              |   |||   | |  |||||  | |||
Db        229 AAGAAGACATCTACCAAGACATCTACT------------------------------- 255

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360
                      ||||||  |  |   ||  | |    |||||||  ||  |  ||
Db        256 ------------ACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420
              |  | ||||  || |||||  ||||       ||||  ||  ||  | |||  | ||  | |||
Db        304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363

Qy        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480
              |||||  |    |||    |   |||||||  |||    ||  ||  |  ||||    ||  |||||
Db        364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423

Qy        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540
              ||  ||||||||||||||  |||||||||  |||||  ||  ||  |||  ||||||    |||
Db        424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483

Qy        541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600
              ||||  |  |||||   |    ||||||    ||  |    |||  |||  ||||||  ||  |||
Db        484 GTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGT 543

Qy        601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660
                 |||||||   |  |  ||  |||||  ||||  ||||||||||  |||  |   ||  ||  ||
Db        544 AACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTC 603

Qy        661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720
              ||  ||||||||||||  |||||  ||  |||  ||||    |  |||||||    ||||||||||  |
Db        604 GCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGC 663

Qy        721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780
              ||||  |||  |  ||  |||||  |||    ||  |       ||  ||||||||||||||||||||
Db        664 TGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTT 723

Qy        781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840
              ||||||||  |||    ||  |  |||||||||          ||||||||  ||||||||||||
Db        724 ACCAACACCGGTGGCGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCC 774

Qy        841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900
              ||||||||  ||||||||||  ||||||  |||||  ||||||||||  |||||||    |||||
Db        775 GGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGC 834

Qy        901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960
              |||||  ||  |||||  ||  |||||     ||||  ||||||||||  |     |||||  |||||
Db        835 TGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCT 894

Qy        961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020
              ||||||||||||||||  |||||||||  |||||||||  |||  ||||||||||  ||||||
Db        895 CTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATG 954

Qy       1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080
              ||||  |||  ||||||||||||||     |  |||  |     ||  ||  ||    ||||||
Db        955 ACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAG 1014

Qy       1081 TAA 1083
              |||
Db       1015 TAA 1017


RESULT 8
```

```
AAL43247
ID   AAL43247 standard; DNA; 1017 BP.
XX
AC   AAL43247;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Rhizopus arrhizus endoglucanase-related coding sequence 4.
XX
KW   Zygomycetes-originated endoglucanase; cellulose binding domain;
KW   fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS   Mucor circinelloides.
XX
PN   WO200242474-A1.
XX
PD   30-MAY-2002.
XX
PF   21-NOV-2001; 2001WO-JP10188.
XX
PR   21-NOV-2000; 2000JP-0354296.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakane A,  Baba Y,  Koga J,  Kubota H;
XX
DR   WPI; 2002-471729/50.
DR   P-PSDB; AAO15055.
XX
PT   Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT   with effect of endoglucanase activity enhanced in processing fibers,
PT   deinking waste paper and improving freeness of paper pulp  -
XX
PS   Disclosure; Page 70-73; 109pp; Japanese.
XX
CC   The invention comprises the amino acid and coding sequences of
CC   zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC   binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC   invention have enhanced endoglucanase activity. The zygomycetes-
CC   originated endoglucanase enzymes of the invention are useful for
CC   processing fibres, de-inking waste paper and improving the freeness of
CC   paper pulp - which is particularly applicable in detergent compositions.
CC   The present DNA sequence represents an endoglucanase-related gene
CC   sequence of the invention.
XX
SQ   Sequence 1017 BP; 233 A; 255 C; 236 G; 293 T; 0 other;
```

Query Match 37.4%; Score 404.6; DB 24; Length 1017;
Best Local Similarity 65.6%; Pred. No. 2.6e-106;
Matches 671; Conservative 0; Mismatches 289; Indels 63; Gaps 3;

```
Qy          61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120
               ||  | |||||   || |||   |  ||    |||||||||| |    |||  ||||
Db          58 GCTGAAGCTGCTTCTTGCAGCTCTGTCTATGGTCAATGTGGTGGCATTGGATGGAGTGGA 117

Qy         121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180
               |||||||| ||||||  ||||||||||||||||||| | |    | |||    | ||||||
Db         118 CCTACCTGTTGTGAAAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCT 177

Qy         181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240
               ||||||  ||||          |    ||  || ||||||  ||  | ||   | || |||
Db         178 CAATGTCTTCC---------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACC 228

Qy         241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300
               |   |||   | |  |||||  | |||
Db         229 AAGAAGACATCTACCAAGACATCTACT------------------------------ 255

Qy         301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360
                       ||||||  | |  |    || | |     |||||||  || | | ||
Db         256 -----------ACCACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAG 303

Qy         361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420
               |  |  ||||  || |||||  ||||      ||||  || || | ||| | ||  | |||

Db         304 ACAACTACCAAGACAACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCT 363

Qy         421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480
               |||||  |    |||   |  |||||| |||   || || | ||||   || |||||
Db         364 TCTTCTGCTGGTTACAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGT 423

Qy         481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540
               || ||||||||||||||  ||||||||| ||||| || || |||  ||||||     |||
Db         424 TATTGGGATTGTTGTAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCT 483

Qy         541 GTTGGCTCCTGTAACAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGT 600
               |||| | |||||  |    ||||||  || |   |||  |||  ||||||  || |||
Db         484 GTTGACACCTGTGCCTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGT 543

Qy         601 GTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTT 660
                  |||||||   | | ||  |||||  |||| |||||||||| ||| |  || || ||
Db         544 AACGGTGGTAATGGTTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTC 603

Qy         661 GCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCC 720
               || |||||||||| |||||  || ||| ||||    | |||||||  ||||||||| |
Db         604 GCTTACGGTTTCGCTGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGC 663

Qy         721 TGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTA 780
               ||||  ||| | || ||||| |||   || |    || ||||||||||||||||||||
Db         664 TGTTATGAATTGACCTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTT 723

Qy         781 ACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCC 840
               |||||||| |||   ||  | |||||||||         |||||||| ||||||||||||
Db         724 ACCAACACCGGTGGCGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCC 774

Qy         841 GGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGT 900
               |||||||| |||||||||| |||||| |||||  |||||||||| |||||||  |||||
Db         775 GGTGGTGGCGTTGGTATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGC 834

Qy         901 TGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCC 960
               ||||| || ||||| || |||||    |||| |||||||||| |    |||||  |||||
Db         835 TGGGGAGCTAGATATGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCT 894

Qy         961 CTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATG 1020
               |||||||||||||||||| |||||||||  |||||||||| ||| |||||||||| ||||||
Db         895 CTTCAAGCTGGTTGTAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATG 954

Qy        1021 ACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAA 1080
               |||| |||  |||||||||||||||    |  | ||| |    || || ||  ||||||
Db         955 ACCTTCAAGGAAGTTACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAG 1014

Qy        1081 TAA 1083
               |||
Db        1015 TAA 1017
```

```
RESULT 9
AAA62730
ID   AAA62730 standard; DNA; 1164 BP.
XX
AC   AAA62730;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 5.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Mucor circinelloides.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
```

XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
DR P-PSDB; AAB09825.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 122-124; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;

Query Match 36.8%; Score 398.6; DB 21; Length 1164;
Best Local Similarity 65.6%; Pred. No. 1.5e-104;
Matches 662; Conservative 0; Mismatches 284; Indels 63; Gaps 3;

```
Qy        75 ATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGA 134
             ||||||    |  ||    ||||| ||||| |     |||  ||||||||||||| |||||
Db       219 ATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGAGTGGACCTACCTGTTGTGA 278

Qy       135 ATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAA 194
             |  ||||||||||||||||| | |     | |||     | ||||||  |||||  ||||
Db       279 AAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAATGTCTTCC--- 335

Qy       195 TGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAA 254
             |    ||  || ||||||  ||  | ||   | || ||| |   |||   | |
Db       336 ------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTAC 389

Qy       255 GAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGAC 314
             |||||  | |||                                          .  ||
Db       390 CAAGACATCTACT----------------------------------------------AC 404

Qy       315 CACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 374
             |||| | |   |   || | |      |||||||  || | | ||   | |  ||||  ||
Db       405 CACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGAC 464

Qy       375 TACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTA 434
             |||||  ||||      ||||  || || | ||| | ||  | ||| ||||| |     ||
Db       465 AACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTA 524

Qy       435 CTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTG 494
             |   |   ||||||| |||   || || | ||||   || ||||| || ||||||||||||
Db       525 CAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTG 584

Qy       495 TAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAA 554
             |||  |||||||||| ||||| || || |||  ||||||    |||||| | |||||
Db       585 TAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGC 644

Qy       555 CAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAG 614
             |    ||||||   || |    |||  |||  ||||||  || |||   |||||||    |
Db       645 CTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGG 704

Qy       615 CTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGC 674
             | ||  |||||  |||| ||||||||||| ||| |  || || || || ||||||||||||
Db       705 TTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGC 764

Qy       675 CGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCAC 734
             |||||  || ||| ||||    | |||||||   |||||||||| |||||  ||| | ||
Db       765 TGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGAC 824

Qy       735 ATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTC 794
             |||||  |||   ||  |    || |||||||||||||||||||| ||||||| |||
Db       825 CTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGG 884

Qy       795 TGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGG 854
             ||  | |||||||||          |||||||| |||||||||||||||||||| |||||
Db       885 CGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGG 935

Qy       855 TATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATA 914
             ||||| |||||| |||||  ||||||||| |||||||  |||||  |||| || |||||
Db       936 TATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATA 995

Qy       915 CGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTG 974
             || |||||    |||| ||||||||||| |    ||||| |||||  ||||||||||||||
Db       996 TGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTG 1055

Qy       975 TAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGT 1034
             ||| |||||||||  |||||||| ||| |||||||||| |||||||||  |||  ||||
Db      1056 TAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGT 1115

Qy      1035 TACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083
             |||||||||    |   | ||| |    || || ||   |||||| |||
Db      1116 TACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1164
```

RESULT 10
AAL43248
ID AAL43248 standard; DNA; 1164 BP.
XX
AC AAL43248;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related coding sequence 5.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Mucor circinelloides.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15056.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 75-78; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.

```
XX
SQ   Sequence 1164 BP; 272 A; 289 C; 266 G; 337 T; 0 other;

  Query Match          36.8%;  Score 398.6;  DB 24;  Length 1164;
  Best Local Similarity   65.6%;  Pred. No. 1.5e-104;
  Matches  662;  Conservative    0;  Mismatches  284;  Indels   63;  Gaps    3;

Qy         75 ATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGA 134
              ||||||    |  ||    ||||| ||||| |      |||  ||||||||||| |||||
Db        219 ATGTAGTTCCGTCTATAGTCAATGCGGTGGCATTGGATGGAGTGGACCTACCTGTTGTGA 278

Qy        135 ATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAA 194
              |  ||||||||||||||||| | |    | |||    | ||||||| ||||||| ||||
Db        279 AAGTGGCTCTACTTGCGTTGCTCAAGAAGGCAACAAATACTACTCTCAATGTCTTCC--- 335

Qy        195 TGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAA 254
                 |    ||  || ||||||  ||  | ||    | || ||| |   |||   | |
Db        336 ------CGGATCCCACAGTAACAATGCTGGTAACGCTAGCAGCACCAAGAAGACATCTAC 389

Qy        255 GAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGAC 314
              |||||  | |||                                          ||
Db        390 CAAGACATCTACT-----------------------------------------AC 404

Qy        315 CACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 374
              |||| | |   |   || | |     |||||||  || | | ||   |  |  |||| ||
Db        405 CACCGCCAAGGCTACTGCTACTGTCACCACCAAGACAGTAACCAAGACAACTACCAAGAC 464

Qy        375 TACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTA 434
              |||||  ||||       ||||  || || | ||| | ||  | ||| ||||| |    ||
Db        465 AACTACCAAGACTAGCACTACTGCCGCTGCTTCTACTTCCACCTCTTCTTCTGCTGGTTA 524

Qy        435 CTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTG 494
              |    |  |||||| |||    || || | ||||   || |||||  || ||||||||||
Db        525 CAAGGTCATCTCTGGCGGTAAATCTGGCAGTGGTTCCACAACTCGTTATTGGGATTGTTG 584

Qy        495 TAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAA 554
              |||  |||||||||| ||||| || || |||  ||||||    |||||| | |||||
Db        585 TAAAGCTTCTTGCAGCTGGCCTGGAAAAGCTTCTGTCACTGGTCCTGTTGACACCTGTGC 644

Qy        555 CAAGGATGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAG 614
              |    ||||||   || |    |||  |||  ||||||  || |||   |||||||    |
Db        645 CTCCAATGGTATCTCTTTATTAGATGCCAATGCTCAAAGTGGTTGTAACGGTGGTAATGG 704

Qy        615 CTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGC 674
              | ||  |||||  |||| |||||||||| ||| |  || || ||  |||||||||||
Db        705 TTTCATGTGTAACAACAACCAACCTTGGGCTGTCAATGATGAGCTCGCTTACGGTTTCGC 764

Qy        675 CGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCAC 734
              |||||  || ||| ||||    | |||||||   |||||||||| |||||  ||| | ||
Db        765 TGCTGCCTCTATTGCTGGCTCCAACGAAGCTGGATGGTGTTGTGGCTGTTATGAATTGAC 824

Qy        735 ATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTC 794
              |||||  |||   || |    || ||||||||||||||||||  |||||||  |||
Db        825 CTTCACTTCTGGCGCTGCTTCTGGAAAGAAGATGGTTGTTCAAGTTACCAACACCGGTGG 884

Qy        795 TGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGG 854
              ||  | ||||||||||                ||||||| ||||||||||||||||||  |||||
Db        885 CGATTTAGGCTCTAAC---------CACTTTGATTTGCAAATGCCCGGTGGTGGCGTTGG 935

Qy        855 TATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATA 914
              |||||  ||||||  |||||  |||||||||| |||||||   |||||  |||||  || |||||
Db        936 TATCTTCAATGGCTGTGCTGCTCAATGGGGCGCTCCCAATGATGGCTGGGGAGCTAGATA 995

Qy        915 CGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTG 974
              || |||||    ||||  |||||||||| |    ||||| |||||  ||||||||||||||
Db        996 TGGTGGTGTCAGCTCTGTCTCTGACTGTGCCTCTCTTCCCTCTGCTCTTCAAGCTGGTTG 1055

Qy        975 TAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGT 1034
              |||  |||||||||  |||||||||  ||| ||||||||||  |||  ||||  ||||
Db       1056 TAAATGGAGATTCAACTGGTTCAAGAACTCTGATAACCCTACCATGACCTTCAAGGAAGT 1115

Qy       1035 TACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083
              |||||||||    |  | ||| |    || || ||   |||||| |||
Db       1116 TACCTGTCCTGCTGAATTAACTACTCGCTCAGGTTGCGAAAGAAAGTAA 1164


RESULT 11
AAA62731
ID   AAA62731 standard; DNA; 1041 BP.
XX
AC   AAA62731;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase nucleotide sequence 6.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; ss.
XX
OS   Phycomyces nitens.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
DR   P-PSDB; AAB09826.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs  -
XX
PS   Claim 44; Page 128-129; 180pp; Japanese.
XX
CC   This sequence encodes an endoglucanase protein. The invention relates
CC   to an endoglucanase of fungal origin which can completely break down
CC   purified cellulose at a concentration of less than 1mg protein/litre,
CC   and produces more than 50% breakdown of cellulose at pH 8.5. The
CC   invention includes endoglucanase protein sequences (see
CC   AAB09825-B09830), endoglucanase nucleotide sequences (see
CC   AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC   identification of the endoglucanase sequences, and in the construction of
CC   vectors containing the polynucleotides. The endoglucanase enzymes are
CC   used for the production of pulp for papermaking and for the production of
CC   animal foodstuffs.
XX
SQ   Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

  Query Match          33.6%;  Score 363.4;  DB 21;  Length 1041;
  Best Local Similarity   61.3%;  Pred. No. 2e-94;
  Matches  652;  Conservative    0;  Mismatches  381;  Indels   30;  Gaps    3;

Qy         21 CTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATGGCCCATGCTGCTGAATGTAG 80
              ||||  |      |  || |||  | ||| | |         |  | ||||||||||| ||
Db          9 CTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAG 68

Qy         81 CAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGAATCTGG 140
              | | |  ||  ||| |||||||| ||||  |||  ||| || ||||||||     || ||
Db         69 CCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGG 128

Qy        141 CTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAATGAAAA 200
              ||  || || || | |       ||||    | ||||| || ||| | ||||| ||  |
Db        129 CTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCA 188

Qy        201 CCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAAGAAGAC 260
                  |          | |   |   | ||| |||||||| ||    |   |||||    |
Db        189 AGT------------CCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGC 236
```

```
Qy        261 TACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 320
Db        237 CACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGC 296

Qy        321 TACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCACTACTAC 380
Db        297 CCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCAC 356

Qy        381 TAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTACTCCGC 440
Db        357 CACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGG---------CTACAGCCC 407

Qy        441 TGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCC 500
Db        408 CATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCC 467

Qy        501 TTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGA 560
Db        468 CTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGA 527

Qy        561 TGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAGCTACAC 620
Db        528 TGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACAT 587

Qy        621 CTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGC 680
Db        588 GTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGC 647

Qy        681 TTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCAC 740
Db        648 CAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCAC 707

Qy        741 CTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCT 800
Db        708 CAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCT 767

Qy        801 TGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTA 860
Db        768 CAGCACCAAC---------CACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTT 818

Qy        861 CAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGG 920
Db        819 CAACGGCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGG 878

Qy        921 TGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTG 980
Db        879 TATTAGCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTG 938

Qy        981 GAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTG 1040
Db        939 GAGATTCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTG 998

Qy       1041 TCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083
Db        999 CCCTGCCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
```

RESULT 12
AAL43249
ID AAL43249 standard; DNA; 1041 BP.
XX
AC AAL43249;
XX
DT 22-AUG-2002 (first entry)
XX
DE Phycomyces nitens endoglucanase-related coding sequence.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Phycomyces nitens.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15057.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Disclosure; Page 81-83; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1041 BP; 225 A; 352 C; 248 G; 216 T; 0 other;

Query Match 33.6%; Score 363.4; DB 24; Length 1041;
Best Local Similarity 61.3%; Pred. No. 2e-94;
Matches 652; Conservative 0; Mismatches 381; Indels 30; Gaps 3;

```
Qy         21 CTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATGGCCCATGCTGCTGAATGTAG 80
Db          9 CTCCATCATCGCTTCCGCCCTTCTCCTCGCTGCCAGCTCCACTTACGCTGCTGAATGCAG 68

Qy         81 CAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGACCTACCTGCTGTGAATCTGG 140
Db         69 CCAAGGCTATGGCCAGTGTGGTGGCAAGATGTGGACTGGTCCCACCTGCTGCACCTCCGG 128

Qy        141 CTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCCCAATGTGTTCCCAATGAAAA 200
Db        129 CTTCACCTGTGTAGGTGCCGAAAACAACGAGTGGTACTCTCAGTGTATCCCCAACGATCA 188

Qy        201 CCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACTACTGAGAGTGCCAAGAAGAC 260
Db        189 AGT------------CCAGGGTAACCCCAAGACCACCACCACCACCACCACCAAGGCTGC 236

Qy        261 TACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCAC 320
Db        237 CACTACCACCAAGGCTCCTGTCACCACCACCAAGGCCACCACCACCACCACCACCAAGGC 296

Qy        321 TACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCTAAGAAGACCACCACTACTAC 380
Db        297 CCCTGTCACCACCACCAAGGCCACTACTACTACCACCACCAAGACCACCACCAAGACCAC 356

Qy        381 TAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCTGCTTCTACAAACTACTCCGC 440
Db        357 CACCACCAAGGCTGCCACCACCACCTCCTCTTCCAACACTGG---------CTACAGCCC 407

Qy        441 TGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCC 500
Db        408 CATTTCTGGTGGCTTCTCTGGAAACGGTCGCACTACCCGCTACTGGGATTGCTGCAAGCC 467

Qy        501 TTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGA 560
Db        468 CTCTTGCGCCTGGGACGGAAAGGCTTCTGTAACTAAGCCTGTACTCACCTGTGCCAAGGA 527

Qy        561 TGGTAAGACTCTTGCTGATAACAACACTCAAAACGGCTGTGTTGGTGGTAGCAGCTACAC 620
```

```
Db        528 TGGTGTCAGCCGTCTCGGTTCCGATGTCCAGAGCGGTTGCGTCGGCGGCCAGGCCTACAT 587

Qy        621 CTGTAATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGC 680

Db        588 GTGCAATGACAACCAGCCCTGGGTTGTCAATGACGACCTTGCCTACGGTTTCGCTGCTGC 647

Qy        681 TTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCAC 740

Db        648 CAGTCTCGGTAGCGCCGGTGCCTCTGCATTCTGCTGCGGCTGTTACGAGCTTACCTTCAC 707

Qy        741 CTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCT 800

Db        708 CAACACTGCTGTCGCTGGCAAGAAGTTTGTCGTCCAGGTCACCAACACCGGTGATGATCT 767

Qy        801 TGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTA 860

Db        768 CAGCACCAAC---------CACTTTGATTTGCAGATGCCCGGCGGTGGTGTCGGCTACTT 818

Qy        861 CAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGG 920

Db        819 CAACGGCTGCCAGTCCCAGTGGAACACCAACACCGATGGCTGGGGTGCTCGCTATGGCGG 878

Qy        921 TGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTG 980

Db        879 TATTAGCTCTATTTCAGAGTGCGACAAGCTTCCTACCCAGTTGCAGGCTGGTTGCAAGTG 938

Qy        981 GAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTG 1040

Db        939 GAGATTCGGATGGTTCAAGAACGCTGACAACCCAGAGGTCACCTTCAAGGCTGTTACTTG 998

Qy       1041 TCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGAAAATAA 1083

Db        999 CCCTGCCGAGATCATTGCCAAGACTGGTTGCGAGCGCAAGTAA 1041
```

RESULT 13
AAA62732
ID AAA62732 standard; DNA; 1043 BP.
XX
AC AAA62732;
XX
DT 25-SEP-2000 (first entry)
XX
DE Endoglucanase nucleotide sequence 7.
XX
KW Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW animal foodstuff; ss.
XX
OS Rhizopus oryzae.
XX
PN WO200024879-A1.
XX
PD 04-MAY-2000.
XX
PF 25-OCT-1999; 99WO-JP05884.
XX
PR 23-OCT-1998; 98JP-0302387.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakamura Y, Moriya T, Baba Y, Yanai K, Sumida N, Nishimura T;
PI Murashima K, Nakane A, Yaguchi T, Koga J, Murakami T, Kono T;
XX
DR WPI; 2000-365117/31.
XX
PT Endoglucanases of fungal origin with high activity under alkaline
PT conditions for production of paper pulp and animal feedstuffs -
XX
PS Claim 44; Page 182-134; 180pp; Japanese.
XX
CC This sequence encodes an endoglucanase protein. The invention relates
CC to an endoglucanase of fungal origin which can completely break down
CC purified cellulose at a concentration of less than 1mg protein/litre,
CC and produces more than 50% breakdown of cellulose at pH 8.5. The
CC invention includes endoglucanase protein sequences (see
CC AAB09825-B09830), endoglucanase nucleotide sequences (see
CC AAA62726-A62732) and primers (AAA62733-A62802) which are used in the
CC identification of the endoglucanase sequences, and in the construction of
CC vectors containing the polynucleotides. The endoglucanase enzymes are
CC used for the production of pulp for papermaking and for the production of
CC animal foodstuffs.
XX
SQ Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;

Query Match 33.5%; Score 362.4; DB 21; Length 1043;
Best Local Similarity 62.9%; Pred. No. 3.8e-94;
Matches 683; Conservative 0; Mismatches 331; Indels 72; Gaps 5;

```
Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60

Db         16 ATGAAGTTCATCACTATCGCCTCCTCCGCCCTCCTTGCCCTCGCCCTTGGCACTGAGATG 75

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Db         76 GCCTCCGCCGCTGAGTGCTCCAAGCTCTACGGACAGTGCGGCGGAAAGAACTGGAACGGC 135

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Db        136 CCCACCTGCTGCGAGAGCGGCTCGACCTGCAAGGTCTCGAATGAC------TACTACAGC 189

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Db        190 CAGTGCCTGCCGAGCG----------------------GCTCCTCGGGAAACAAGTCG 225

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Db        226 AGCGAGTCGGCCCACAAGAAGACCACGACCGCTGCCCACAAGAAGACCACGAC------- 278

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Db        279 -------------------------CGCCGCTCACAAGAAGACTACGACCGCTCCCGCC 312

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Db        313 AAGAAGACCACGACCGTCGCCAAGGCTTCGACTCCGTCCAACTCGAGCAGCTCGTCTTCG 372

Qy        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Db        373 G------GAAAGTACAGCGCTGTCAGCGGTGGCGCTAGCGGCAACGGCGTCACTACCCGC 426

Qy        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Db        427 TACTGGGACTGCTGCAAGGCTTCGTGCTCGTGGCCCGGCAAGGCTAACGTCAGCTCGCCT 486

Qy        541 GTTGGCTCCTGTAACAAGGATGG---TAAGACTCTTGCTGATAACAACACTCAAAACGGC 597

Db        487 GTCAAGTCCTGCAACAAGGACGGCGTCACCGCTCTTAGCGACTCCAACGCCCAGTCCGGC 546

Qy        598 TGTGTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGAC 657

Db        547 TGCAACGGCGGCAACTCCTACATGTGCAACGACAACCAGCCATGGGCTGTCAACGACAAC 606

Qy        658 CTTGCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGT 717

Db        607 CTTGCTTACGGTTTCGCTGCCGCTGCCATTAGCGGCGGTGGCGAGAGCCGCTGGTGCTGC 666

Qy        718 GCCTGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAA 777

Db        667 TCCTGCTTCGAGCTCACCTTCACCTCCACCAGCGTTGCTGGCAAGAAGATGGTCGTCCAG 726

Qy        778 GTAACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATG 837

Db        727 GTCACCAACACTGGCGGTGACCTTGGCAGCTCGACCGGTGCCCACTTCGATCTCCAGATG 786

Qy        838 CCCGGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGAT 897
```

```
Db        787 CCCGGCGGCGGCGTCGGCATCTTCAACGGATGCTCGTCCCAGTGGGGCGCTCCCAACGAC 846

Qy        898 GGTTGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCT 957

Db        847 GGCTGGGGCTCGCGCTACGGCGGCATCAGCTCCGCCAGCGACTGCTCGTCCCTCCCCAGC 906

Qy        958 GCCCTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACC 1017

Db        907 GCCCTCCAGGCCGGCTGCAAGTGGCGCTTCAACTGGTTCAAGAACGCCGACAACCCGTCC 966

Qy       1018 ATGACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGA 1077

Db        967 ATGACCTACAAGGAGGTCACCTGCCCCAAGGAGATCACCGCTAAGACCGGATGCTCGCGC 1026

Qy       1078 AAATAA 1083

Db       1027 AAGTAA 1032
```

RESULT 14
AAL43250
ID AAL43250 standard; DNA; 1043 BP.
XX
AC AAL43250;
XX
DT 22-AUG-2002 (first entry)
XX
DE Rhizopus arrhizus endoglucanase-related codon-optimised DNA sequence.
XX
KW Zygomycetes-originated endoglucanase; cellulose binding domain;
KW fibre processing; waste paper de-inking; paper pulp; ds; gene.
XX
OS Rhizopus arrhizus.
OS Synthetic.
XX
PN WO200242474-A1.
XX
PD 30-MAY-2002.
XX
PF 21-NOV-2001; 2001WO-JP10188.
XX
PR 21-NOV-2000; 2000JP-0354296.
XX
PA (MEIJ) MEIJI SEIKA KAISHA LTD.
XX
PI Nakane A, Baba Y, Koga J, Kubota H;
XX
DR WPI; 2002-471729/50.
DR P-PSDB; AAO15052.
XX
PT Cellulose-binding domain-lacking Zygomycetes-originated endoglucanase,
PT with effect of endoglucanase activity enhanced in processing fibers,
PT deinking waste paper and improving freeness of paper pulp -
XX
PS Example 10; Page 84-86; 109pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequences of
CC zygomycetes-originated endoglucanase enzymes lacking the cellulose
CC binding domain. The zygomycetes-originated endoglucanase enzymes of the
CC invention have enhanced endoglucanase activity. The zygomycetes-
CC originated endoglucanase enzymes of the invention are useful for
CC processing fibres, de-inking waste paper and improving the freeness of
CC paper pulp - which is particularly applicable in detergent compositions.
CC The present DNA sequence represents an endoglucanase-related gene
CC sequence of the invention.
XX
SQ Sequence 1043 BP; 212 A; 370 C; 291 G; 170 T; 0 other;

Query Match 33.5%; Score 362.4; DB 24; Length 1043;
Best Local Similarity 62.9%; Pred. No. 3.8e-94;
Matches 683; Conservative 0; Mismatches 331; Indels 72; Gaps 5;

```
Qy          1 ATGAAGTTCCTTACCATTGCCTCCTCCGCTATCTTGGCACTTGCCGTCGGTACTGAAATG 60

Db         16 ATGAAGTTCATCACTATCGCCTCCTCCGCCCTCCTTGCCCTCGCCCTTGGCACTGAGATG 75

Qy         61 GCCCATGCTGCTGAATGTAGCAAGGCTTACTACCAATGTGGTGGTAAGAACTGGGATGGA 120

Db         76 GCCTCCGCCGCTGAGTGCTCCAAGCTCTACGGACAGTGCGGCGGAAAGAACTGGAACGGC 135

Qy        121 CCTACCTGCTGTGAATCTGGCTCTACTTGCGTTGATTATCCTGACAATCCTTTCTACTCC 180

Db        136 CCCACCTGCTGCGAGAGCGGCTCGACCTGCAAGGTCTCGAATGAC------TACTACAGC 189

Qy        181 CAATGTGTTCCCAATGAAAACCTCACCTCCACTAACAAATCTTCTCACAAAACCACCACT 240

Db        190 CAGTGCCTGCCGAGCG----------------------GCTCCTCGGGAAACAAGTCG 225

Qy        241 ACTGAGAGTGCCAAGAAGACTACCACTACTAAAGGTTCCAAGAAGACCACCACTACTGAA 300

Db        226 AGCGAGTCGGCCCACAAGAAGACCACGACCGCTGCCCACAAGAAGACCACGAC------- 278

Qy        301 GCCTCTAAGAAGACCACCACTACTGAAGCTTCCAAGAAGACCACCACTACTGAAGCCTCT 360

Db        279 -------------------------CGCCGCTCACAAGAAGACTACGACCGCTCCCGCC 312

Qy        361 AAGAAGACCACCACTACTACTAAGAAGGCTTCTACCTCCACTTCCTCTTCCTCTTCTTCT 420

Db        313 AAGAAGACCACGACCGTCGCCAAGGCTTCGACTCCGTCCAACTCGAGCAGCTCGTCTTCG 372

Qy        421 GCTTCTACAAACTACTCCGCTGTCTCTGGTGGTGCCTCCGGTAATGGTGAAACCACTCGC 480

Db        373 G------GAAAGTACAGCGCTGTCAGCGGTGGCGCTAGCGGCAACGGCGTCACTACCCGC 426

Qy        481 TACTGGGATTGTTGTAAGCCTTCTTGCAGTTGGCCCGGTAAGGCTGATGTCACCTCCCCT 540

Db        427 TACTGGGACTGCTGCAAGGCTTCGTGCTCGTGGCCCGGCAAGGCTAACGTCAGCTCGCCT 486

Qy        541 GTTGGCTCCTGTAACAAGGATGG---TAAGACTCTTGCTGATAACAACACTCAAAACGGC 597

Db        487 GTCAAGTCCTGCAACAAGGACGGCGTCACCGCTCTTAGCGACTCCAACGCCCAGTCCGGC 546

Qy        598 TGTGTTGGTGGTAGCAGCTACACCTGTAATGACAATCAACCTTGGGTTGTTAGCGACGAC 657

Db        547 TGCAACGGCGGCAACTCCTACATGTGCAACGACAACCAGCCATGGGCTGTCAACGACAAC 606

Qy        658 CTTGCCTACGGTTTCGCCGCTGCTTCCATTTCTGGTGGTAGCGAAGCTACTTGGTGTTGT 717

Db        607 CTTGCTTACGGTTTCGCTGCCGCTGCCATTAGCGGCGGTGGCGAGAGCCGCTGGTGCTGC 666

Qy        718 GCCTGTTTCGAACTCACATTCACCTCTACTGCCGTCAAGGGTAAGAAGATGGTTGTTCAA 777

Db        667 TCCTGCTTCGAGCTCACCTTCACCTCCACCAGCGTTGCTGGCAAGAAGATGGTCGTCCAG 726

Qy        778 GTAACCAACACTGGTTCTGACCTTGGCTCTAACACTGGTGCTCACTTTGACTTGCAAATG 837

Db        727 GTCACCAACACTGGCGGTGACCTTGGCAGCTCGACCGGTGCCCACTTCGATCTCCAGATG 786

Qy        838 CCCGGTGGTGGTGTTGGTATCTACAATGGTTGTGCCACTCAATGGGGTGCTCCCACCGAT 897

Db        787 CCCGGCGGCGGCGTCGGCATCTTCAACGGATGCTCGTCCCAGTGGGGCGCTCCCAACGAC 846

Qy        898 GGTTGGGGTGCAAGATACGGCGGTGTTTCTTCTGCCTCTGACTGTTCTAACCTTCCTTCT 957

Db        847 GGCTGGGGCTCGCGCTACGGCGGCATCAGCTCCGCCAGCGACTGCTCGTCCCTCCCCAGC 906

Qy        958 GCCCTTCAAGCTGGTTGTAAGTGGAGATTCGGCTGGTTCAAAAACGCTGATAACCCAACC 1017

Db        907 GCCCTCCAGGCCGGCTGCAAGTGGCGCTTCAACTGGTTCAAGAACGCCGACAACCCGTCC 966

Qy       1018 ATGACCTACAAACAAGTTACCTGTCCCAAGGCTATCACTGCCAAGTCTGGCTGTTCAAGA 1077

Db        967 ATGACCTACAAGGAGGTCACCTGCCCCAAGGAGATCACCGCTAAGACCGGATGCTCGCGC 1026

Qy       1078 AAATAA 1083
```

```
Db       1027 AAGTAA 1032


RESULT 15
AAQ14857
ID   AAQ14857 standard; DNA; 1473 BP.
XX
AC   AAQ14857;
XX
DT   18-FEB-1992  (first entry)
XX
DE   Fusarium oxysporum DSM 2672 endoglucanase coding sequence.
XX
KW   cellulase; cellulose; ss.
XX
OS   Fusarium oxysporum.
XX
FH   Key             Location/Qualifiers
FT   CDS             97..1227
FT                   /*tag=  a
XX
PN   WO9117243-A.
XX
PD   14-NOV-1991.
XX
PF   08-MAY-1991;   91WO-DK00123.
XX
PR   22-APR-1991;   91DK-0000736.
PR   09-MAY-1990;   90DK-0001159.
XX
PA   (NOVO ) NOVO NORDISK A/S.
XX
PI   Rasmussen G,  Mikkelsen JM,  Schulein M,  Patkar SA,  Hagen F;
PI   Hjort CM,  Hastrup S;
XX
DR   WPI; 1991-353765/48.
DR   P-PSDB; AAR15272.
XX
PT   Cellulase prepn. comprising endoglucanase enzyme - used in
PT   detergents for cellulose-contg. fabrics or to improve drainage of
PT   paper pulp
XX
PS   Claim 11; Page 52; 67pp; English.
XX
CC   The cellulase (i.e. endoglucanase) gene was isolated from a
CC   F.oxysporum cDNA library by screening with probes based on the
CC   H.insolens 43kD endoglucanase sequence. Positive clones were
CC   subjected to PCR amplification using 43kD-specific oligonucleotides
CC   as primers. The amplified DNA was sequenced.
CC   See also AAQ14856.
XX
SQ   Sequence 1473 BP; 343 A; 453 C; 337 G; 340 T; 0 other;

  Query Match             20.1%;  Score 218.2;  DB 12;  Length 1473;
  Best Local Similarity   63.0%;  Pred. No. 1.8e-52;
  Matches  394;  Conservative    0;  Mismatches  213;  Indels   18;  Gaps    3;

Qy        449 GTGGTGCCTCCGGTAATGGTGAAACCACTCGCTACTGGGATTGTTGTAAGCCTTCTTGCA 508
Db        143 GTGCTGCTTCTGGAAGCGGTCACTCTACTCGATACTGGGATTGCTGCAAGCCTTCTTGCT 202

Qy        509 GTTGGCCCGGTAAGGCTGATGTCACCTCCCCTGTTGGCTCCTGTAACAAGGATGGTAAGA 568
Db        203 CTTGGAGCGGAAAGGCTGCTGTCAACGCCCCTGCTTTAACTTGTGATAAGAACGACAACC 262

Qy        569 CTCTTGCTGATAACAACACTCAAAACGGCTGT---GTTGGTGGTAGCAGCTACACCTGTA 625
Db        263 CCATTTCCAACACCAATGCTGTCAACGGTTGTGAGGGTGGTGGTTCTGCTTATGCTTGCA 322

Qy        626 ATGACAATCAACCTTGGGTTGTTAGCGACGACCTTGCCTACGGTTTCGCCGCTGCTTCCA 685
Db        323 CCAACTACTCTCCCTGGGCTGTCAACGATGAGCTTGCCTACGGTTTCGCTGCTACCAAGA 382

Qy        686 TTTCTGGTGGTAGCGAAGCTACTTGGTGTTGTGCCTGTTTCGAACTCACATTCACCTCTA 745
Db        383 TCTCCGGTGGCTCCGAGGCCAGCTGGTGCTGTGCTTGCTATGCTTTGACCTTCACCACTG 442

Qy        746 CTGCCGTCAAGGGTAAGAAGATGGTTGTTCAAGTAACCAACACTGGTTCTGACCTTGGCT 805
Db        443 GCCCCGTCAAGGGCAAGAAGATGATCGTCCAGTCCACCAACACTGGAGGTGATCTCGGCG 502

Qy        806 CTAACACTGGTGCTCACTTTGACTTGCAAATGCCCGGTGGTGGTGTTGGTATCTACAATG 865
Db        503 ACAAC---------CACTTCGATCTCATGATGCCCGGCGGTGGTGTCGGTATCTTCGACG 553

Qy        866 GTTGTGCCACTCAATGGGGTGCTCCCACCGATGGTTGGGGTGCAAGATACGGCGGTGTTT 925
Db        554 GCTGCACCTCTGAGTTCGG------CAAGGCTCTCGGCGGTGCCCAGTACGGCGGTATCT 607

Qy        926 CTTCTGCCTCTGACTGTTCTAACCTTCCTTCTGCCCTTCAAGCTGGTTGTAAGTGGAGAT 985
Db        608 CCTCCCGAAGCGAATGTGATAGCTACCCCGAGCTTCTCAAGGACGGTTGCCACTGGCGAT 667

Qy        986 TCGGCTGGTTCAAAAACGCTGATAACCCAACCATGACCTACAAACAAGTTACCTGTCCCA 1045
Db        668 TCGACTGGTTCGAGAACGCCGACAACCCTGACTTCACCTTTGAGCAGGTTCAGTGCCCCA 727

Qy       1046 AGGCTATCACTGCCAAGTCTGGCTG 1070
Db        728 AGGCTCTCCTCGACATCAGTGGATG 752


Search completed: June 17, 2003, 11:49:07
Job time : 243.225 secs
```